LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-218088

# FY04 LDRD Final Report

## Protein-Protein Integration Mapping of the Human DNA Damage Response Pathway

J. S. Albala

January 11, 2006

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

# FY04 LDRD Final Report
# Protein-Protein Integration Mapping of the Human DNA Damage Response Pathway
# LDRD Project Tracking Code: 04-FS-014
**Joanna S. Albala, Principal Investigator**

*Research Goal:*

This LDRD supported a development leave for Dr. Joanna Albala to participate in the laboratory of Dr. Marc Vidal at the Dana Farber Cancer Institute of Harvard University. Dr. Vidal is an expert in high-throughput cloning and high-throughput yeast two-hybrid automated analysis for comparative proteomics. The research goal was to develop protein interaction (interactome) maps for relevant microbes to apply to GTL efforts in the BBRP program. During this time, Dr. Albala aquired the techniques for high-throughput recombination-based cDNA cloning using the Gateway system and has applied these methods for generating recombinant proteins in *E. coli* from the Human genome sequence (Rual et al., 2004). In addition, Dr. Albala has continued to develop and expand upon methods for high-throughput cloning of human cDNAs into a miniaturized baculovirus system for heterologous protein production. Methods for miniaturized co-immunoprecipitation have been devised using this technique whereby protein-protein interactions can be quickly screened in a miniaturized, *in vivo* format. These technologies will be applied to the human interactome map that the Vidal laboratory is embarking upon whereby a 30,000 x 30,000 human gene matrix will be screened to analyze for protein-protein interactions within the human proteome.

*Specific Aims:*

1). To gain expertise in automated, high-throughput cloning and yeast two-hybrid analysis for interactome mapping for human interactome mapping in collaboration with the efforts and goals of the Vidal Laboratory.

2). To integrate expertise in high-throughput miniaturized baculoviral expression for protein production/co-immunoprecipitation techniques in baculovirus in the Vidal Laboratory to validate protein interactions.

*Accomplishments:*

The first version of the human ORFeome has been derived using an improved Gateway recombinational cloning approach. Using cDNA

clones originally derived from the LLNL-IMAGE collection, 8,187 human open reading frames (ORFs), representing at least 7,354 human genes, have been placed into this system as mini-pools of PCR-amplified products. This is the human ORFeome version 1.1 (hORFeome v1.1) collection. After assessing the overall quality of this collection, the hORFeome v1.1 was used for heterologous protein expression in two different expression systems at proteome scale. The hORFeome v1.1 represents a central resource for the cloning of large sets of human ORFs in various settings for functional proteomics of many types, and will serve as the foundation for the mapping of the Human Interactome.

Having the hORFeome collection in hand allowed for the rapid cloning of thousands of ORFs into multiple protein expression vectors. These clones were used as a platform for proteome-wide protein expression and analysis and the hORFeome v1.1 was used for high-throughput protein expression. 282 ORF mini-pools were transferred by automated methods from the Entry constructs into Destination vectors for two different protein expression systems: i) in *E. coli* as amino-terminal His6-tag fusion (hexa-histidine tag), and ii) in mammalian cells as a carboxy-terminal GFP (green fluorescent protein) fusion.

For protein expression in *E. coli*, 282 randomly selected Entry constructs were cloned via an LR reaction into pDEST17 vector (Invitrogen) for expression in *E. coli* (Braun et al. 2002). The resulting products were transformed into *E. coli* BL21 Star (DE)pLysS strain (Invitrogen), and heterologous protein expression was induced with 1 mM IPTG. Cells were lysed in SDS-PAGE sample dye followed by polyacrylamide gel electrophoresis. Proteins were transferred to PVDF membranes and standard Western blotting was performed. Recombinant proteins expressed from the pDEST17 vector incorporate an N-terminal His6 tag, which was detected by immunoblotting first with the anti-polyhistidine His-1 mouse monoclonal primary antibody (Sigma) at 1:2000 dilution, and then with a goat anti-mouse HRP conjugated secondary antibody (Calbiochem) at 1:1000 dilution. Visualization was performed using enhanced chemiluminescence (Amersham).

Bacterial protein expression was assessed by Western blot analysis using an antibody against His6, observing a band of the appropriate size 55% of the time (see Figure 1). The rate of successful protein expression decreased with increasing ORF length, as is commonly observed with recombinant expression from bacteria, with the set of short lengths showing 79% successful expression of proteins in the 30-35 kDa range, the medium set showing 57% successful expression of proteins in the 55-69 kDa

range, and the long set showing 30% successful expression of proteins in the 90-114 kDa range (Fig. 1). These results compare favorably with previous proteome-scale expression of human proteins in bacteria (Braun et al. 2002)

The same sets of proteins expressed in 293T mammalian cells produced a detectable signal 68% of the time (71%, 79% and 54% success for short, medium and long ORF sets respectively) (Rual et al., 2004). Most of the failures of expression in mammalian cells likely arise from poor transfection. Overall, the high success rate obtained for protein expression suggests that the hORFeome v1.1 is a useful resource for the expression of proteins at proteome scale in multiple protein expression systems. Thus the Human ORFeome will be a valuable resource for reverse proteomics applications such as protein-protein interaction mapping projects forthcoming from this project (Li et al. 2004).

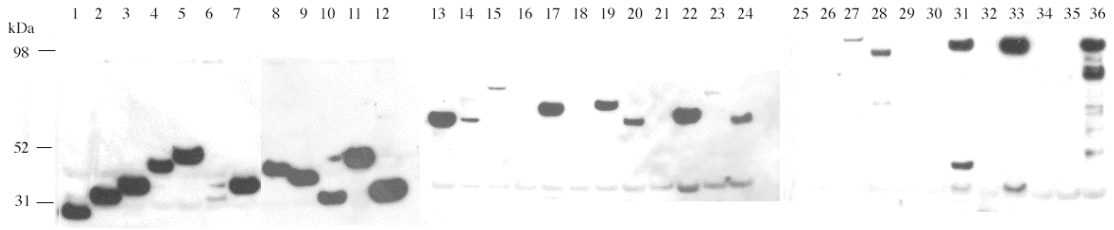


**Figure 1:**
Recombinant protein expression in *E. coli*. The 282 Entry constructs were transferred into a His6 N-terminal fusion vector (pDEST 17). Representative samples from the 282 small (Lanes 1-12), medium (Lanes 13-24), and long ORFs (Lanes 25-36) are shown. The positions of molecular weight markers (31-98 kDa) are indicated. All visible proteins migrate at the expected size.

*References:*

Braun P, LaBaer J. (2003). High throughput protein production for functional proteomics.

Trends Biotechnol. 2003 Sep;21(9):383-8

Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual JF, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, Van Den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Harper JW, Cusick ME, Roth FP, Hill DE, **Vidal M.** (2004). A map of the interactome network of the metazoan C. elegans. *Science*. Jan 23;303(5657):540-3.

Rual JF, Hirozane-Kishikawa T, Hao T, Bertin N, Li S, Dricot A, Li N, Rosenberg J, Lamesch P, Vidalain PO, Clingingsmith TR, Hartley JL, Esposito D, Cheo D, Moore T, Simmons B, Sequerra R, Bosak S, Doucette-Stamm L, Le Peuch C, Vandenhaute J, Cusick ME, Albala JS, Hill DE, Vidal M. (2004). Human ORFeome version 1.1: a platform for reverse proteomics.
Genome Res.14(10B):2128-35.